



# Min-max certainty equivalence principle and differential games

Pierre Bernhard, Alain Rapaport

## ► To cite this version:

Pierre Bernhard, Alain Rapaport. Min-max certainty equivalence principle and differential games. [Research Report] RR-2019, INRIA. 1993. inria-00074652

**HAL Id: inria-00074652**

**<https://inria.hal.science/inria-00074652>**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Coll Dif



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

***Min-Max Certainty  
Equivalence Principle  
and Differential Games***

Pierre BERNHARD  
Alain RAPAPORT

N° 2019  
Septembre 1993

PROGRAMME 5

Traitement du signal,  
automatique et  
productique

**R**apport  
de recherche

1993

# Min-Max Certainty Equivalence Principle and Differential Games

Pierre Bernhard and Alain Rapaport

INRIA Sophia-Antipolis

July 1993

## Abstract

This paper presents a version of the Certainty Equivalence Principle, usable for nonlinear, variable end-time, partial observation Zero-Sum Differential Games, which states that under the unicity of the solution to the auxiliary problem, optimal controllers can be derived from the solution of the related perfect observation game. An example is provided where in one region, the new extended result holds, giving an optimal control, and in another region, the unicity condition is not met, leading indeed to a non-certainty equivalent optimal controller.

# Principe d'Equivalence à la Certitude Min-Max et Jeux Différentiels

## Résumé

Cet article propose une extension du Principe d'Equivalence à la Certitude, dédiée aux jeux différentiels à somme nulle, non linéaires, à horizon variable et partiellement observés par l'un des joueurs. Les contrôleurs optimaux sont obtenus à partir de la résolution du jeu en information parfaite, et à condition qu'un certain problème auxiliaire admette une solution unique. Un exemple, où dans une région de l'espace d'états le principe s'applique, et dans une autre un contrôleur non équivalent à la certitude est optimal, est présenté.

# 1 Introduction

It has already been pointed out by Isaacs in 1965 (chapter 12 of [Isa65]) that differential games with incomplete information are worth being studied, but a general theory seemed to be out of the scope of the simple calculus of variations. For players perfectly knowing the state, the optimal controllers exhibited by the Isaacs/Breakwell theory are pure state-feedbacks. When at least one of the players is denied the complete knowledge of the state, the history of past observations and controls is extremely relevant and, as a consequence, the structure of optimal controllers may be radically different. Since then, the literature has rather been concerned with linear games and convex information sets (see for instance the chapter XV of [KS77]), or stochastic modelizations where typically, disturbances are modeled as gaussian variables, and the criterion is replaced by an expectation. This leads to ‘stochastic control’. But, very few progresses have been attempted in the deterministic non linear min-max area.

The certainty equivalence principle, originally discovered by Bernhard in 1990 [Ber90a], was motivated by  $H^\infty$  robust control problems. In the state space approach, the game theoretic view point leads to an easy determination of the equations of the  $H^\infty$ -optimal controllers via the well-known Riccati equations, considering the disturbance as another player’s control (‘worst-case design’). For this purpose, the principle is used with linear quadratic fixed terminal time problems, and [BB91] was the first book completely devoted to this approach. It has already been pointed out by Bernhard that the principle holds for non-linear problems [Ber90a]. Nevertheless, two difficulties were met for using it with Isaacs’ differential games. The proof needed the surjective behavior of the observation map:  $\forall x, v \mapsto y = h(x, v)$  must be onto, which does not allow simple observed outputs such as  $y = h(x)$ . Didinsky, Başar and Bernhard [DBB93] have proposed an extension where the unicity of the worst disturbance is replaced by the unicity of the worst state, on the condition that the set of compatible states is open, and some extra regularity of the ‘cost to come’ function is met. Terminal time should again be fixed, and so here also, we cannot use these results in the scope of Isaacs’ differential games. This paper proposes a generalization of the principle overcoming these difficulties. It can be used for a larger class of classical zero-sum differential games (where terminal time is defined as the first entrance inside a target), where one of the players has a partial knowledge of the state, even if the set of compatible states is not open, but on the condition that no barrier exists in the related perfect information game. Nevertheless, for the qualitative study of a game where a barrier exists, we show that the principle holds with a somewhat different criterion, if the boundary of the target is smooth enough. In this paper, we also exhibit an example, which shows that when the unicity of the auxiliary problem is not met, the value of the game can be far from verifying Isaacs’ dynamic programming equation.

## 2 The perfect information problem

Let us consider a zero-sum differential game, as in [Isa65] defined by the following components:

- A dynamical system

$$S(t_0, x_0, u, v) \begin{cases} \frac{dx}{dt} &= f(t, x, u, v) \\ x(t_0) &= x_0 \end{cases}$$

where  $t \in [t_0, +\infty)$ ,  $x(t) \in X = \mathbb{R}^n$ ,  $x_0 \in X_0 \subset X$ ,  $u(t) \in U$ , closed subset in  $\mathbb{R}^p$ , and  $v(t) \in V$ , closed subset in  $\mathbb{R}^q$ . Admissible open loop controls will be the sets  $\mathcal{U}$  and  $\mathcal{V}$  of measurable functions from  $[t_0, +\infty)$  into  $U$  and  $V$  respectively.

- A capture set  
 $\mathcal{T}$  is an open subset of  $\mathbb{R} \times X$ .
- A criterion

$$J(t_0, x_0, u, v) = M(t_f, x(t_f)) + \int_{t_0}^{t_f} L(t, x(t), u(t), v(t))dt + N(x_0).$$

We will consider two kinds of terminal time  $t_f$ . When  $(t_0, x_0) \notin \mathcal{T}$  and  $x(\cdot)$  is solution of  $S(t_0, x_0, u, v)$ ,

$$\begin{aligned} t_f^1(t_0, x_0, u, v) &= \inf_{t > t_0} \{(t, x(t)) \in \mathcal{T}\}, \\ t_f^2(t_0, x_0, u, v) &= \text{Arg min}_{t \geq t_0} \left\{ M(t, x(t)) + \int_{t_0}^t L(\tau, x(\tau), u(\tau), v(\tau))d\tau + N(x_0) \right\}, \end{aligned}$$

and will note the criterion  $J^i$  if  $t_f^i$  is used ( $i \in \{1, 2\}$ ).

Typically, in pursuit-evasion games, one may be interested in the qualitative or quantitative problem (game of kind or game of degree). If the capture occurs, the criterion associated with  $M = 0$ ,  $L = 1$  and  $t_f = t_f^1$  gives the minimal interception time. If the target is described by the sign of a function  $C(\cdot)$ :  $\mathcal{T} = \{(t, x) \in \mathbb{R} \times X \mid C(t, x) < 0\}$ , then the sign of the criterion associated with  $M = C$ ,  $L = 0$  and  $t_f = t_f^2$  provides the qualitative result: capture or evasion. Other mixed criterions may be of interest.

- Sets of admissible controllers

We will consider classes  $\Phi$  and  $\Psi$  of admissible feedback strategies  $u(t) = \phi(t, x(t))$  and  $v(t) = \psi(t, x(t), u(t))$  satisfying the following properties:

- Open-loop controls are admissible:  $\mathcal{U} \subset \Phi$  and  $\mathcal{V} \subset \Psi$ .

- $\Phi$  and  $\Psi$  are closed by concatenation.
- $\forall(\phi, \psi) \in \Phi \times \Psi$ , the system (S) admits an unique solution  $x(\cdot)$ , leading to admissible controls  $u(\cdot) = \phi(\cdot, x(\cdot)) \in \mathcal{U}$  and  $v(\cdot) = \psi(\cdot, x(\cdot), u(\cdot)) \in \mathcal{V}$ .

Thus,  $\psi(\cdot)$  is a discriminating feedback. It is known, though, that if ‘Isaacs’ condition’ is met, the optimal  $\psi^*$  is in fact a simple state feedback. These properties do not uniquely specify the pair  $(\Phi, \Psi)$ , but we know that such pairs exist, and the above assumptions suffice to justify Isaacs’ equation (see [Ber87]).

According to the kind of terminal time, let us now introduce the following problems:

**Problem  $\mathcal{P}^i(t_0, X_0)$  ( $i \in \{1, 2\}$ )**

If the initial condition  $(t_0, x_0) \in (\{t_0\} \times X_0) \setminus \mathcal{T}$ , does a Isaacs’ Value function:

$$V^i(t_0, x_0) = \min_{\phi \in \Phi} \max_{\psi \in \Psi} J^i(t_0, x_0, \phi, \psi) - N(x_0)$$

exist?

For the following, we will need some assumptions to bound the terminal time:

- For problem  $\mathcal{P}^1$ , the system will be assumed to be capturable over  $(t_0, X_0)$ , in the very strong sense that:

$$\exists T^1(t_0, X_0) < +\infty, \exists u \in \mathcal{U}, \forall x_0 \in X_0, \forall v \in \mathcal{V}, t_f^1(t_0, x_0, u, v) < T^1(t_0, X_0).$$

This is satisfied if, for instance,  $\exists T^1$  such that  $[T^1, +\infty) \times X \subset \mathcal{T}$ .

- For problem  $\mathcal{P}^2$ , we will assume that:

$$\exists T^2(t_0, X_0) < +\infty, \forall x_0 \in X_0, \forall (u, v) \in \mathcal{U} \times \mathcal{V}, \exists t_f^2(t_0, x_0, u, v) < T^2(t_0, X_0).$$

Let us stress that problem  $\mathcal{P}^2$  reads:

$$\min_{u(\cdot)} \max_{v(\cdot)} \min_{t_f} \left[ M(t_f, x(t_f)) + \int_{t_0}^{t_f} L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \right].$$

Let us also recall the classical Hamilton-Jacobi-Isaacs theorem [Ber87].

**Proposition 1** *If there exists a  $C^1$  function  $V$  over  $([t_0, T^1(t_0, X_0)] \times X) \setminus \mathcal{T}$ , solution of the partial differential equation:*

$$\begin{cases} -\frac{\partial V}{\partial t} &= \min_{u \in \mathcal{U}} \max_{v \in \mathcal{V}} H(t, x, \frac{\partial V}{\partial x}, u, v) \\ V(t, x) &= M(t, x) \text{ if } (t, x) \in \partial \mathcal{T} \end{cases}$$

where  $H(t, x, \lambda, u, v) = L(t, x, u, v) + \lambda^t f(t, x, u, v)$  is the Hamiltonian of the system, then  $V$  is a solution of the problem  $\mathcal{P}^1(t_0, X_0)$ . Moreover, any admissible  $(\phi^*, \psi^*)$  such that:

$$\begin{cases} \phi^*(t, x) & \in \text{Arg} \min_{u \in U} H(t, x, \frac{\partial V}{\partial x}, u, \psi^*(t, x, u)) \\ \psi^*(t, x, u) & \in \text{Arg} \max_{v \in V} H(t, x, \frac{\partial V}{\partial x}, u, v) \end{cases}$$

are optimal state feedbacks.

### Proof

- Let us consider a controller  $\phi(.) \in \Phi$ , and  $u(.)$  be its open loop representation against  $\psi^*$ . Then, from Isaacs' equation:

$$\forall t, \quad \frac{\partial V}{\partial t}(t, x(t)) + \frac{\partial V}{\partial x}(t, x(t)) \cdot f(t, x(t), u(t), \psi^*(t, x(t), u(t))) + L(t, x(t), u(t), \psi^*(t, x(t), u(t))) \geq 0.$$

Thus,

$$\forall t > t_0, \quad V(t, x(t)) - V(t_0, x_0) + \int_{t_0}^t L(\tau, x(\tau), u(\tau), \psi^*(\tau, x(\tau), u(\tau))) d\tau \geq 0,$$

where  $x(.) \in S(t_0, x_0, u, \psi^*)$ .

At  $t_f = t_f^1(t_0, x_0, u, \psi^*)$ , we conclude:

$$\begin{aligned} V(t_0, x_0) + N(x_0) &\leq M(t_f, x(t_f)) + \int_{t_0}^{t_f} L(\tau, x(\tau), u(\tau), \psi^*(\tau, x(\tau), u(\tau))) d\tau + N(x_0) \\ &= J^1(t_0, x_0, \phi, \psi^*). \end{aligned}$$

- Changing the role of  $\phi$  and  $\psi$ , we obtain symmetrically the other saddle-point like inequality:

$$V(t_0, x_0) + N(x_0) \geq J^1(t_0, x_0, \phi^*, \psi), \quad \forall \psi(.) \in \Psi$$

□

A similar version of this proposition for the problem  $\mathcal{P}^2$  is:

**Proposition 2** *If there exists a  $C^1$  function  $V$  over  $([t_0, T^2(t_0, X_0)) \times X) \setminus \mathcal{T}$  such that:*  
 $\forall (t, x) \in ([t_0, T^2(t_0, X_0)) \times X) \setminus \mathcal{T},$

$$V(t, x) \leq M(t, x),$$

$$\min_{u \in U} \max_{v \in V} \left\{ \frac{\partial V}{\partial t} + H(t, x, \frac{\partial V}{\partial x}, u, v) \right\} \geq 0,$$

$$V(t, x) \neq M(t, x) \implies \min_{u \in U} \max_{v \in V} \left\{ \frac{\partial V}{\partial t} + H(t, x, \frac{\partial V}{\partial x}, u, v) \right\} = 0,$$

$$\forall (\phi, \psi) \in \Phi \times \Psi, \exists \hat{t} < +\infty \mid V(\hat{t}, x(\hat{t})) = M(\hat{t}, x(\hat{t}))$$

when  $x(.) \in S(t_0, x_0, \phi, \psi)$ ,

$H(t, x, \lambda, u, v) = L(t, x, u, v) + \lambda^t f(t, x, u, v)$  is the Hamiltonian of the system.

Then  $V$  is a solution of the problem  $\mathcal{P}^2(t_0, X_0)$ . Moreover, any admissible controller satisfying:

$$\begin{cases} \phi^*(t, x) \in \text{Arg} \min_{u \in U} H(t, x, \frac{\partial V}{\partial x}, u, \psi^*(t, x, u)) \\ \psi^*(t, x) \in \text{Arg} \max_{v \in V} H(t, x, \frac{\partial V}{\partial x}, u, v) \end{cases}$$

are optimal state feedbacks, and an optimal stopping time is:

$$t^* = \min \{t \geq t_0 \mid V(t, x^*(t)) = M(t, x^*(t))\}$$

if  $x^*(.) \in S(t_0, x_0, \phi^*, \psi^*)$ .

### Proof

Let  $(t_0, x_0)$  be an initial condition,

- Let  $\phi(.) \in \Phi$  be a u-controller,  $(u(.), v^*(.))$  the open-loop representation of the pair  $(\phi, \psi^*)$ , and  $x(.)$  the related trajectory when the game is played along with  $(\phi, \psi^*)$ . Then for all  $t \geq t_0$ , we have:

$$\begin{aligned} & \frac{\partial V}{\partial t}(t, x(t)) + \frac{\partial V}{\partial t}(t, x(t)) \cdot f(t, x(t), u(t), v^*(t)) + L(t, x(t), u(t), v^*(t)) \geq 0 \\ \implies & \forall t \geq t_0, V(t, x(t)) - V(t_0, x_0) + \int_{t_0}^t L(\tau, x(\tau), u(\tau), v^*(\tau)) d\tau \geq 0 \\ \implies & \forall t \geq t_0, M(t, x(t)) + \int_{t_0}^t L(\tau, x(\tau), u(\tau), v^*(\tau)) d\tau \geq V(t_0, x_0) \end{aligned}$$

and so:

$$V(t_0, x_0) + N(x_0) \leq J^2(t_0, x_0, u, v^*).$$

- Let  $\psi(.) \in \Psi$  be a v-controller, and  $u^*(.), v(.), x(.)$  defined in a similar manner as previously,



– If  $V(t_0, x_0) = M(t_0, x_0)$ , it is clear that:

$$V(t_0, x_0) + N(x_0) \geq J^2(t_0, x_0, u^*, v).$$

– If  $V(t_0, x_0) \neq M(t_0, x_0)$ , then as long as  $V(t, x(t)) \neq M(t, x(t))$  we get:

$$\begin{aligned} \frac{\partial V}{\partial t}(t, x(t)) + \frac{\partial V}{\partial t}(t, x(t)) \cdot f(t, x(t), u^*(t), v(t)) + L(t, x(t), u^*(t), v(t)) &\leq 0 \\ \implies V(t, x(t)) + \int_{t_0}^t L(\tau, x(\tau), u^*(\tau), v(\tau)) d\tau &\leq V(t_0, x_0). \end{aligned}$$

At  $\bar{t} = \min \{t \mid V(t, x(t)) = M(t, x(t))\}$ , we get:

$$V(t_0, x_0) \geq M(\bar{t}, x(\bar{t})) + \int_{t_0}^{\bar{t}} L(\tau, x(\tau), u^*(\tau), v(\tau)) d\tau$$

and so, a fortiori

$$V(t_0, x_0) + N(x_0) \geq J^2(t_0, x_0, u^*, v).$$

With these two inequalities, we conclude that  $V(t_0, x_0)$  is the value of the game when a pair of controllers  $(\phi^*, \psi^*)$  min-maximizing the Hamiltonian is played, and that  $t^*$  is an optimal terminal time.

□

We can now propose a generalization of the notion of the target for both problems  $\mathcal{P}^1$  and  $\mathcal{P}^2$ :

**Definition** Let  $\mathcal{T}^1 = \mathcal{T}$  and  $\mathcal{T}^2 = \{(t, x) \mid V(t, x) = M(t, x)\}$ .

### 3 An imperfect information structure

For the following, we will not write the indices of  $J$ ,  $V$ ,  $t_f$  and  $T$  when a definition is shared by both problems  $\mathcal{P}^1$  and  $\mathcal{P}^2$ .

For a given pair  $(t_0, X_0)$ , let us write  $T = T(t_0, X_0)$ ,  $\mathcal{U} = \mathcal{U}|_{[t_0, T]}$ ,  $\mathcal{V} = \mathcal{V}|_{[t_0, T]}$ ,  $\omega = (x_0, v) \in \Omega = X_0 \times \mathcal{V}$ ,  $S(t_0, u, \omega) = S(t_0, x_0, u, v)$  and  $J(t_0, u, \omega) = J(t_0, x_0, u, v)$ .

Instead of a perfect knowledge of the state, let us consider an observation scheme for the player  $u$ , which to each pair  $(u, \omega) \in \mathcal{U} \times \Omega$  makes correspond a family of subsets  $\Omega_{(\cdot)} \in \mathcal{F}$ :

$$\left\{ \begin{array}{ll} \Omega_{(\cdot)} : [t_0, T] &\longrightarrow \mathcal{P}(\Omega) \\ t &\longmapsto \Omega_t \in \mathcal{F}_t \subset \mathcal{P}(\Omega) \end{array} \right.$$

such that  $u$  knows, at time  $t$ , that  $x_0$  and the control of its opponent  $v(\cdot)$  belong to the subset  $\Omega_t$  (We shall give more concrete examples shortly).

Let us also introduce truncated information sets:

$$\begin{cases} u^t &= \{\tau \mapsto u(\tau), \tau \leq t\} \in \mathcal{U}^t, \\ v^t &= \{\tau \mapsto v(\tau), \tau \leq t\} \in \mathcal{V}^t, \\ \omega^t &= (x_0, v^t) \in \Omega^t, \end{cases}$$

and define:

$$\Omega_t^t = \{\omega^t \mid \omega \in \Omega_t\}.$$

### Hypotheses H1

**Hypothesis H1a** The process  $\Omega_{(\cdot)}$  is consistent:

$$\forall t \in [t_0, T], \omega \in \Omega_t.$$

**Hypothesis H1b** The process  $\Omega_{(\cdot)}$  is perfect recall:

$$\forall t' > t, \Omega_{t'} \subset \Omega_t.$$

**Hypothesis H1c** The process  $\Omega_{(\cdot)}$  is nonanticipative:

$$\omega^t \in \Omega_t^t \iff \omega \in \Omega_t.$$

Typically, an observed output:  $y(t) = h(t, x(t), v(t))$  leads to such a knowledge. Thus,

$$\Omega_t = \{(x_0, v) \in \Omega \mid \text{for } x(\cdot) \in S(t_0, x_0, u, v), \forall \tau \leq t, y(\tau) = h(\tau, x(\tau), v(\tau))\}$$

is the equivalence class of all disturbances  $\omega$  that together with  $u^t$  lead to the same observed output history up to time  $t$ :  $y^t$ .

We shall need the following *restricted observation sets* for the problem  $\mathcal{P}^i$ :

$$\tilde{\Omega}_t = \Omega_t \cap \{\omega \mid \forall \tau \leq t, x(\tau) \notin T^i\}.$$

**Proposition 3**  $\tilde{\Omega}_t$  satisfies Hypothesis H1 up to  $t_f^i$ .

### Proof

Consistence up to  $t_f^i$  is by definition, the decreasing character is preserved by intersection of two decreasing subset families, and the same holds true for the non anticipative character.  $\square$

As a matter of fact, in the sequel we shall only use  $\tilde{\Omega}_t$  and thus omit the  $\sim$ , meaning that we have included the restriction into the definition of  $\Omega_t$ .

Finally, let us define a class  $\mathcal{M}_{[t_0, T]}$  of admissible controllers

$$\left\{ \begin{array}{ll} \mu : [t_0, T] \times \mathcal{F}_t & \longrightarrow \mathcal{U} \\ (t, \Omega_t) & \longmapsto \mu(t, \Omega_t) \in U \end{array} \right.$$

such that  $\forall \omega \in \Omega$ , the pair  $(\mu, \omega)$  generates a well defined unique trajectory. Remark that due to the definition of  $\Omega_t$ , it is indeed a class of nonanticipative controllers.

## 4 The imperfect information problem

### 4.1 The problem

**Problem  $\mathcal{P}_I^i(t_0, X_0)$**  ( $i \in \{1, 2\}$ )

If for every  $x_0 \in X_0$ ,  $(t_0, x_0) \notin \mathcal{T}$ , does there exist a controller  $\mu$  attaining:

$$\min_{\mu \in \mathcal{M}_{[t_0, T]}} \max_{\omega \in \Omega} J^i(t_0, \mu, \omega)$$

and if yes, how to build one?

### 4.2 The auxiliary problem

- The auxiliary function  $G$

**Hypothesis H2** The problem  $\mathcal{P}^i(t_0, X_0)$  admits a saddle-point  $(\phi^*, \psi^*)$  and a value function  $V^i$  of class  $C^1$ .

Under this hypothesis, for all admissible  $(u, \omega) \in \mathcal{U} \times \Omega$ , we define:

$$G_t(u, \omega) = G_t(u^t, \omega^t) = V^i(t, x(t)) + \int_{t_0}^t L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0)$$

where  $x(\cdot)$  is solution of  $S(t_0, x_0, u, v)$ .

- The auxiliary problem  $Q_t(u^t, \Omega_t^t)$

$$\max_{\omega^t \in \Omega_t^t} G_t(u^t, \omega^t).$$

We shall write  $\widehat{\omega}_t = (\widehat{x}_0, \widehat{v}_t)$  for a maximizing disturbance and  $\widehat{x}_t(\cdot)$  for the induced trajectory.

- The controller  $\hat{\mu}$

**Hypothesis H3a** For any  $(u, \Omega_{(\cdot)}) \in \mathcal{U} \times \mathcal{F}$ , and any  $t \in [t_0, T^i]$ , the state at time  $t$   $\hat{x}_t(t)$  is unique among all the state solutions  $S(t_0, \hat{x}_{0t}, u, \hat{v}_t)$  at time  $t$ .  $\hat{x}_t(t)$  represents then the unique ‘worst state’ at time  $t$ , compatible with the past observations, and will be simply written  $\hat{x}(t)$ .

We wish to define  $\hat{\mu}$  as  $\hat{\mu}(t, \Omega_t) = \phi^*(t, \hat{x}(t))$ . As a matter of fact,  $\hat{x}(t)$  also depends on past  $u$ ’s (i.e on  $u^t$ ), therefore the above definition is actually a fixed point, that can be made explicit as follows.

Let us define for any  $\tau \in [t_0, t]$ ,  $u_t^*(\tau) = \phi^*(\tau, \hat{x}_t(\tau))$  where  $\phi^*(\cdot)$  is the  $u$ -optimal feedback for the problem  $(\mathcal{P}^i)$ , and  $\hat{x}_t(\cdot)$  is the solution, restricted to  $[t_0, t]$ , of  $S(t_0, \hat{x}_{0t}, u, \hat{v}_t)$ . Then consider the causal application  $\mathcal{A}_t: (u^t, \Omega_t^i) \mapsto u_t^*(t)$ . This defines in a natural way an application  $\mathcal{A}: \mathcal{U} \times \mathcal{F} \rightarrow \mathcal{U}$ . The strategy  $\hat{\mu}(\cdot, \Omega_{(\cdot)})$  will be defined as a fixed point of the application  $\mathcal{A}$ , for any  $t \geq t_0$  as long as  $\{\hat{x}(\tau) \notin \mathcal{T}^i, \forall \tau \in [t_0, t]\}$ .

**Hypothesis H3b**  $\hat{\mu}$  is uniquely defined and admissible.

Computing  $\bar{u}(\cdot)$ , knowing the solution of the game in perfect information and getting  $\Omega_{(\cdot)}$  consists in solving, for each time  $t$ , an auxiliary problem  $Q_t(\bar{u}^t, \Omega_t^i)$ , which gives the value of  $\bar{u}$  at time  $t$ , if an unique worst state is met. Choosing  $\hat{\mu}(\cdot)$  amounts to playing as if the real state of the system were  $\hat{x}_t(t)$ . This is the reason why we shall call  $\hat{\mu}$  a ‘certainty equivalent’ controller.

## 5 The Min-Max Certainty Equivalence Principle

Here again, we will not distinguish between problems  $\mathcal{P}_I^1$  and  $\mathcal{P}_I^2$  when parts of the proof are common.

### 5.1 The principle

**Proposition 4** *Under the hypotheses H1-H3,  $\hat{\mu}(\cdot)$  is an optimal controller and it guarantees:*

$$\max_{\omega \in \Omega} J(t_0, \hat{\mu}, \omega) = \max_{x_0 \in X_0} [V(t_0, x_0) + N(x_0)] = \Gamma(t_0).$$

#### Proof

For a given  $\bar{\omega} = (\bar{x}_0, \bar{v}) \in \Omega$ , let us call  $\bar{u}(\cdot)$  the open-loop representation of the controller  $\hat{\mu}$ ,  $t_f$  the corresponding final time, and  $\Omega_{(\cdot)}$  the information process. Any other  $\omega \in \Omega_t$  would produce the same  $\bar{u}^t$ . Then, let us define, for any  $t$  in  $[t_0, t_f]$ ,

$$G(t, \omega) = G_t(\bar{u}^t, \omega^t).$$

Notice that although  $G$  is defined as a function from  $[t_0, t_f] \times \Omega$  into  $\mathbb{R}$ , it actually only depends upon the restriction  $\omega^t$  of  $\omega$  to  $[t_0, t]$ . Thus, using also hypothesis H1c, we get

$$\Gamma(t) = \max_{\omega \in \Omega_t} G(t, \omega) = \max_{\omega^t \in \Omega_t^t} G_t(\bar{u}^t, \omega^t).$$

Let us also notice that  $\Gamma$  is well defined, because  $\Omega_t$  is never empty, thanks to the hypothesis H1a.

We are now going to prove that the function  $\Gamma(\cdot)$  is nonincreasing over  $[t_0, t_f]$ . To each time instant  $t$ , associate the function:

$$\Gamma_t(\tau) : \tau \longmapsto \max_{\omega \in \Omega_t} G(\tau, \omega).$$

where the constraint bearing upon  $\omega$  has been frozen at  $\omega \in \Omega_t$ . For  $t$  in  $(t_0, t_f)$  and  $h > 0$  such that  $t + h \leq t_f$  and  $t - h \geq t_0$ , we have, according to the hypothesis H1b:

$$\begin{cases} \Gamma(t - h) \leq \Gamma_t(t - h), \\ \Gamma(t) = \Gamma_t(t), \\ \Gamma(t + h) \geq \Gamma_t(t + h). \end{cases}$$

So, the upper Dini derivative of  $\Gamma$  verifies:

$$D^+ \Gamma(t) = \limsup_{t' \rightarrow t} \frac{\Gamma(t') - \Gamma(t)}{t' - t} \leq \limsup_{t' \rightarrow t} \frac{\Gamma_t(t') - \Gamma_t(t)}{t' - t}.$$

Thanks to Danskin's theorem [Ber90b],  $\Gamma_t(\cdot)$  has left and right derivatives  $\Gamma_t^l(\cdot)$  and  $\Gamma_t^r(\cdot)$ :

$$\Gamma_t^l(\tau) = \min_{\hat{\omega} \in \hat{\Omega}_t} \frac{\partial G(\tau, \hat{\omega})}{\partial \tau}, \quad \Gamma_t^r(\tau) = \max_{\hat{\omega} \in \hat{\Omega}_t} \frac{\partial G(\tau, \hat{\omega})}{\partial \tau}.$$

When  $\tau = t$ , these last expressions depend only on the restriction of  $\hat{\Omega}_t$  to the interval  $[t_0, t]$ . For each  $\hat{\omega}^t = (\hat{x}_0, \hat{v}^t)$  realizing the maximum of the problem  $Q_t(\bar{u}^t, \Omega_t^t)$ , the generated state  $\hat{x}(\cdot) \in S(t_0, \hat{x}_0, u, \hat{v})$  at time  $t$  is the same, thanks to the hypothesis H3a. Let call it  $\hat{x}(t)$ . For any maximizing  $\hat{\omega}^t$ , we have:

$$\begin{aligned} \frac{\partial G(t, \hat{\omega})}{\partial t} &= \frac{\partial V}{\partial t}(t, \hat{x}(t)) + \frac{\partial V}{\partial x}(t, \hat{x}(t)) \cdot f(t, \hat{x}(t), \bar{u}(t), \hat{v}(t)) + L(t, \hat{x}(t), \bar{u}(t), \hat{v}(t)) \\ &= \frac{\partial V}{\partial t}(t, \hat{x}(t)) + H \left( t, \hat{x}(t), \frac{\partial V}{\partial x}(t, \hat{x}(t)), \bar{u}(t), \hat{v}(t) \right). \end{aligned}$$

For the problem  $\mathcal{P}_I^1$ , according to Isaacs' equation and the choice of  $\hat{\mu}$ , this last expression is non positive. For the problem  $\mathcal{P}_I^2$ , the Isaacs' equation also holds thanks to the definition of

$\tilde{\Omega}_{(\cdot)}$ . One concludes that the Dini derivative of  $\Gamma$  is non positive, and therefore  $\Gamma$  is decreasing (see Appendix). So,  $\bar{x}$  being the solution of  $S(t_0, \bar{u}, \bar{\omega})$ , and by H1a,  $\bar{\omega} \in \Omega_{(\cdot)}$ , thus we have:

$$J(t_0, \hat{\mu}, \bar{\omega}) = M(t_f, \bar{x}(t_f)) + \int_{t_0}^{t_f} L(\tau, \bar{x}(\tau), \bar{u}(\tau), \bar{v}(\tau)) d\tau + N(x_0) \leq \Gamma(t_f) \leq \Gamma(t_0)$$

This bound being uniform in  $\bar{\omega}$ , we get:

$$\max_{\omega \in \Omega} J(t_0, \hat{\mu}, \omega) \leq \Gamma(t_0).$$

Conversely, for any controller  $\mu(\cdot)$ , a possible  $\omega$  is

$$\omega^* = \left( x_0^* = \text{Arg} \max_{x_0 \in X_0} V(t_0, x_0), v^*(\cdot) = \psi^*(\cdot, x^*(\cdot), \mu(\cdot)) \right)$$

along the trajectory  $x^*(\cdot)$  and then, according to the other saddle-point inequality of Isaacs' equation:

$$\max_{\omega \in \Omega} J(t_0, \mu, \omega) \geq J(t_0, \mu, \omega^*) = \Gamma(t_0).$$

Hence,  $\hat{\mu}$  is a min-max controller, and even a saddle point controller, and the theorem is proved.

□

We recall also the following result:

**Proposition 5** *If  $\forall (u, \Omega_{(\cdot)}) \in \mathcal{U} \times \mathcal{F}$ , there exists  $\tau^* \in (t_0, T)$  such that  $\sup_{\omega \in \Omega_{\tau^*}} G_{\tau^*}(u, \omega) = +\infty$ , then  $\forall \mu \in \mathcal{M}, \sup_{\omega \in \Omega} J(u, \omega) = +\infty$ .*

**Proof**

For the proof, we refer to [Ber90a], [BB91].

□

This is important to notice because it is to get this result that we constrain  $\omega$  to  $\Omega_t$ .

## 6 The case without unicity

When the auxiliary problem admits an unique solution and the principle holds, the value of the game is the same as in the perfect information case, and so Isaacs' equation is also verified. For the general case, it seems to be reasonable to look after a kind of dynamic

programming equation, where the maximization in  $v$  may be replaced by a maximization in  $\omega$ . Unfortunately, only one inequality can be proved, and we give further an example that shows that we may not expect something better. When the certainty equivalence principle does not hold, none of the trajectories compatible with the observation may be an extremal of the problem in perfect information.

## 6.1 A dynamic programming inequality

Assume there exists an optimal controller  $\hat{\mu}$  solution of the problem  $\mathcal{P}_I^i(t_0, X_0)$  when  $N(\cdot) = 0$ . We shall call  $U(t_0, X_0)$  its value function. For any  $\omega \in \Omega$ , let  $\Omega_{(\cdot)}$  be the information process observed when  $u$  plays  $u^\omega(\cdot) = \hat{\mu}(\cdot, \Omega_{(\cdot)})$ . Then, let us write  $x^\omega(\cdot)$  the solution of the system  $S(t_0, u^\omega, \omega)$ . We claim the following property:

**Proposition 6** *If  $\forall t \in [t_0, T^i]$ ,  $\forall \omega^t \in \Omega^t$ , the problem  $\mathcal{P}_I^i(t, \{x^{\omega^t}(t)\})$  admits a solution  $U(t, \{x^{\omega^t}(t)\})$ , then:*

$$\forall t \in (t_0, T^i), U(t_0, X_0) \geq \max_{\omega^t \in \Omega^t} \left\{ \int_{t_0}^t L(\tau, x^{\omega^t}(\tau), u^{\omega^t}(\tau), v^{\omega^t}(\tau)) d\tau + U(t, \{x^{\omega^t}(t)\}) \right\}$$

### Proof

Let  $\bar{\omega}^t = (x_0, v^{\bar{\omega}^t})$  be an element of  $\Omega^t$ , for any  $v \in \mathcal{V}|_{[t, T^i]}$ , define  $\omega = \bar{\omega}^t.v = (x_0, v^{\bar{\omega}^t}.v)$  (where  $\bar{v}.v$  defines the concatenation of  $\bar{v}$  and  $v$ ).

$$\begin{aligned} U(t_0, X_0) &\geq \max_{v \in \mathcal{V}|_{[t, T^i]}} J^i(t_0, \hat{\mu}, \bar{\omega}^t.v) \\ &= \max_{v \in \mathcal{V}|_{[t, T^i]}} \left[ \int_{t_0}^t L(\tau, x^{\bar{\omega}^t}(\tau), u^{\bar{\omega}^t}(\tau), v^{\bar{\omega}^t}(\tau)) d\tau + J^i(t, x^{\bar{\omega}^t}(t), \hat{\mu}, v) \right]. \end{aligned}$$

Thanks to the hypothesis H1c of the nonanticipative behavior of the observation process  $\Omega_{(\cdot)}$ , the first term does no longer depend on the future disturbance. So,

$$\begin{aligned} U(t_0, X_0) &\geq \int_{t_0}^t L(\tau, x^{\bar{\omega}^t}(\tau), u^{\bar{\omega}^t}(\tau), v^{\bar{\omega}^t}(\tau)) d\tau + \max_{v \in \mathcal{V}|_{[t, T^i]}} J^i(t, x^{\bar{\omega}^t}(t), \hat{\mu}, v) \\ &\geq \int_{t_0}^t L(\tau, x^{\bar{\omega}^t}(\tau), u^{\bar{\omega}^t}(\tau), v^{\bar{\omega}^t}(\tau)) d\tau + U(t, \{x^{\bar{\omega}^t}(t)\}) \end{aligned}$$

□

## 6.2 An example

Let us consider the following problem in the plane:

$$\begin{cases} \dot{x} = -cu + b \cos v \\ \dot{y} = -a + b \sin v \end{cases} \quad \begin{matrix} a > b > c > 0 \\ u \in [-1, 1] \end{matrix}, \quad v \in [0, 2\pi)$$

with the end-point condition:

$$y(0) > 0 \quad , \quad y(t_f) = 0,$$

and the objective:

$$\min_{u \in \mathcal{U}} \max_{v \in \mathcal{V}} |x(t_f)|.$$

### 6.2.1 The perfect information solution

When  $\Phi$  and  $\Psi$  are admissible state feedback classes, one can easily derive a saddle point solution thanks to Isaacs' differential game theory.

Let us write the Isaacs' equation:

$$(I) \begin{cases} b\sqrt{\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2} - c\left|\frac{\partial V}{\partial x}\right| - a\frac{\partial V}{\partial y} = 0 \\ V(x, 0) = |x|, \end{cases}$$

and check that:

$$V(x, y) = |x| + r^+ y,$$

where  $r^+$  is the positive root of

$$b\sqrt{1 + r^2} - c - ar = 0$$

is a solution of (I) in the viscosity sense.

It is immediate to verify that  $V$  is a  $C^1$  solution, except on the  $x = 0$  axis, where sub and superdifferentials are:

$$\begin{cases} \partial^- V(0, y) = \emptyset \\ \partial^+ V(0, y) = [-1, 1] \times \{r^+\} \end{cases}$$

and we have to check only one viscosity inequality:

$$\forall \rho \in [0, 1] \quad , \quad H^*(\rho, r^+) = b\sqrt{\rho^2 + (r^+)^2} - c\rho - ar^+ \leq 0$$

$$\begin{cases} H(0, r^+) = (b - a)r^+ < 0 \\ H(\rho, r^+) = \rho \left[ b\sqrt{1 + \left(\frac{r^+}{\rho}\right)^2} - c - a\frac{r^+}{\rho} \right] \leq 0 \end{cases}$$

So,

$$V(x, y) = |x| + \frac{-ac + \sqrt{(ac)^2 + (a^2 - b^2)(b^2 - c^2)}}{a^2 - b^2} y$$



is the value function of the problem.

Notice that when  $x_0 \neq 0$ , there exist optimal (saddle-point) open loop strategies:

- If  $x_0 > 0$ ,

$$\begin{cases} u^* &= +1 \\ (\sin v^*, \cos v^*) &= \left( \frac{r^+}{\sqrt{1 + (r^+)^2}}, \frac{1}{\sqrt{1 + (r^+)^2}} \right) \end{cases}$$

- If  $x_0 < 0$ ,

$$\begin{cases} u^* &= -1 \\ (\sin v^*, \cos v^*) &= \left( \frac{r^+}{\sqrt{1 + (r^+)^2}}, \frac{-1}{\sqrt{1 + (r^+)^2}} \right) \end{cases}$$

If the initial state stands at  $x_0 = 0$ , two symmetric trajectories are possible. Whatever is the choice of  $u, v$  can force the state to definitely go to one side of the  $y$  axis: It is a dispersal line for the maximizer.

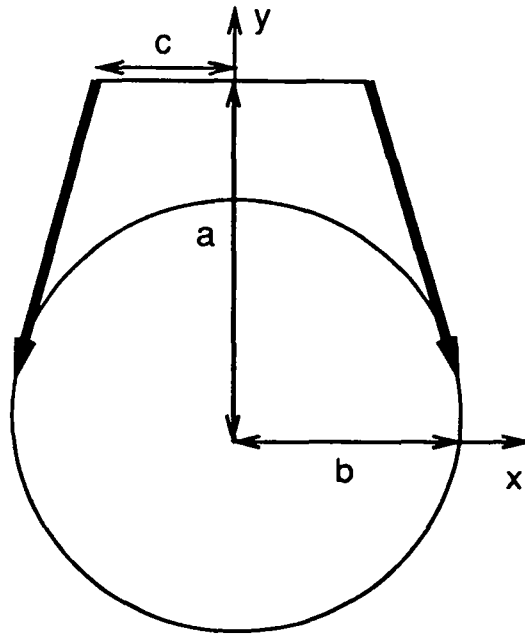


Figure 1: Vectogram and semi-permeable directions

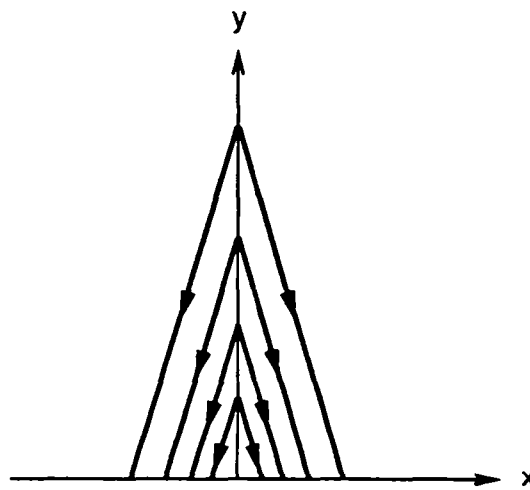


Figure 2: Extremals field

### 6.2.2 An imperfect information scheme

Assume now that the player  $u$  has only the knowledge of  $x(0) = x_0$  and  $\{y(\cdot)\}_{[t_0, t]}$  at time  $t$ . Here, the information scheme is:

$$\begin{cases} X_0 &= \{(x_0, y_0)\} \\ \Omega_t &= \left\{ ((x_0, y_0), v(\cdot)) \mid \forall \tau \in [0, t], \sin v(\tau) = \frac{\dot{y}(\tau) + a}{b} \right\} \end{cases}$$

We are looking for an optimal controller  $\mu(x_0, \Omega_{(\cdot)})$  against the worst  $v(\cdot)$ .

**Proposition 7** *The open-loop controller:*

$$\begin{cases} u^*(t) &= \operatorname{sgn}(x_0) & 0 \leq t \leq \frac{|x_0|}{c} \\ u^*(t) &= 0 & t > \frac{|x_0|}{c} \end{cases}$$

*is optimal.*

**Proof**

Let  $v^*$  be a disturbance realizing  $\max_{v(\cdot)}$  against  $u^*$ . Then the final time is defined as:

$$\int_0^{t_f} (a - b \sin v^*(\tau)) d\tau = y_0.$$

Two cases have to be considered:

- $t_f < \frac{|x_0|}{c}$

Then,

$$\begin{aligned} x(t_f) &= x_0 - \operatorname{sgn}(x_0)ct_f + b \int_0^{t_f} \cos v^*(\tau) d\tau \\ &= \operatorname{sgn}(x_0) [|x_0| - ct_f] + b \int_0^{t_f} \cos v^*(\tau) d\tau. \end{aligned}$$

So,

$$\max_{v(\cdot)} |x(t_f)| = \max_{v(\cdot)} \operatorname{sgn}(x_0)x(t_f).$$

The control  $v^*$  is necessarily solution of this last one player optimal control problem. Thanks to the open-loop saddle point solution of the game in perfect information:  $(u_{PI}^*, v_{PI}^*)$ , and the fact that  $u_{PI}^* = u^*|_{[0, t_f]}$ ,  $v^* = v_{PI}^*$  is optimal.

So,  $J(u^*, v^*) = J(u_{PI}^*, v_{PI}^*)$  which is optimal, corroborating the certainty equivalence principle.

Notice that the current theory accounts for that fact, not any previous one: beyond the variable end time, which is of little consequence here, the theory of [Ber90a] does not apply because the observation function is not surjective in  $v$ , and that of [DBB93] does not either because  $\widehat{X}$  is not open.

- $t_f \geq \frac{|x_0|}{c}$

Let  $\mu(\cdot)$  be a controller, and consider the open-loop control  $\bar{v}^*(\cdot) = \pi - v^*(t)$ . It leads to the same final time  $t_f$  and the same function of time  $y(\cdot)$  than  $v^*(\cdot)$ , when it is played against the same  $\mu(\cdot)$  (i.e. it always belong to the same set  $\Omega_t$  as  $v^*$ ). Let us write  $u(\cdot)$  the open-loop representation of  $\mu(\cdot)$  played along with  $v^*(\cdot)$  or  $\bar{v}^*(\cdot)$ .

$$x(t_f) = x_0 - c \int_0^{t_f} u(\tau) d\tau + b \int_0^{t_f} \cos v(\tau) d\tau.$$

where  $v$  may be either  $v^*$  or  $\bar{v}^*$ . So,

$$\begin{aligned} \max_{v(\cdot)} J(\mu(\cdot), v(\cdot)) &\geq \max \left\{ \left| x_0 - c \int_0^{t_f} u(\tau) d\tau + b \int_0^{t_f} \cos v^*(\tau) d\tau \right|, \right. \\ &\quad \left. \left| x_0 - c \int_0^{t_f} u(\tau) d\tau - b \int_0^{t_f} \cos v^*(\tau) d\tau \right| \right\} \\ &= \left| x_0 - c \int_0^{t_f} u(\tau) d\tau \right| + b \left| \int_0^{t_f} \cos v^*(\tau) d\tau \right| \\ &\geq b \left| \int_0^{t_f} \cos v^*(\tau) d\tau \right| = J(u^*, v^*) = \max_v J(\hat{\mu}, v) \end{aligned}$$

and the result is proved.

□

### 6.2.3 Calculation of the value

Two cases are possible:

- $u^*(\cdot)$  is constant until  $t_f$

Then, the value is the same as in the perfect information scheme:

$$J_1(x_0, y_0) = |x_0| + \frac{-ac + \sqrt{(ac)^2 + (a^2 - b^2)(b^2 - c^2)}}{a^2 - b^2} y_0$$

on the condition that:

$$t_{f_1}(y_0) = \frac{y_0}{a - b \sin v_1^*} = \frac{y_0}{a - b \frac{r^+}{\sqrt{1+(r^+)^2}}} < \frac{|x_0|}{c}.$$

- $u^*(.)$  switches to 0 before  $t_f$

Determining the worst  $v(.)$  against  $u^*(.)$  is a one-player optimal control problem, with a non-autonomous non-continuous differential system. Let us transform it into an autonomous one, introducing the variable  $m(.)$ :

$$m(0) = x_0, \dot{m} = -cu.$$

The system can then be re-written, for a positive  $x_0$ :

$$f_- : \begin{cases} \dot{x} &= b \cos v - c \\ \dot{y} &= b \sin v - a \\ \dot{m} &= -c \end{cases} \quad f_+ : \begin{cases} \dot{x} &= b \cos v \\ \dot{y} &= b \sin v - a \\ \dot{m} &= 0 \end{cases}$$

with the commutation surface  $\mathcal{S} : \{m = 0\}$  from  $f_-$  towards  $f_+$ . Classically,

$$\begin{aligned} H_-^* &= b\sqrt{\lambda_x^2 + \lambda_y^2} - c\lambda_x - a\lambda_y - c\lambda_m, \\ H_+^* &= b\sqrt{\lambda_x^2 + \lambda_y^2} - a\lambda_y \end{aligned}$$

The transversality condition at the crossing of  $\mathcal{S}$  imposes only a discontinuity of  $\lambda_m$ . So, the optimal  $v_2^*$ :

$$\begin{cases} \cos v_2^* &= \frac{\lambda_x}{\sqrt{\lambda_x^2 + \lambda_y^2}} \\ \sin v_2^* &= \frac{\lambda_y}{\sqrt{\lambda_x^2 + \lambda_y^2}} \end{cases}$$

is constant along an extremal. The final condition on the target imposing  $|\lambda_x| = 1$ ,  $\lambda_y = \frac{b}{\sqrt{a^2 + b^2}}$  is the solution of  $H_+^* = 0$ :

$$\begin{cases} |\cos v_2^*| &= \frac{\sqrt{a^2 - b^2}}{a} \\ \sin v_2^* &= \frac{b}{a} \end{cases}$$

from which one can deduce:

$$\begin{cases} t_f &= \frac{y_0}{a - b \sin v_2^*} = \frac{a}{a^2 - b^2} y_0 \\ J_2(x_0, y_0) &= b \left| \int_0^{t_f} \cos v_2^* d\tau \right| = \frac{b}{\sqrt{a^2 - b^2}} y_0 \end{cases}$$

on the condition that:

$$t_{f_2}(y_0) = \frac{y_0}{a - b \sin v_2^*} \geq \frac{|x_0|}{c}.$$

It is easy to check that  $\sin v_1^* < \sin v_2^*$ , and so  $t_{f_1} < t_{f_2}$ . When  $|x_0| \in [ct_{f_1}, ct_{f_2}]$ , both hypotheses are valid. The worst  $v(\cdot)$  is the one which leads to the greater cost. The switching line is defined by  $J_1(x_0, y_0) = J_2(x_0, y_0)$ .

As a conclusion, the value function verifies the:

**Proposition 8**

$$\begin{cases} U(x_0, y_0) = |x_0| + \frac{-ac + \sqrt{(ac)^2 + (a^2 - b^2)(b^2 - c^2)}}{a^2 - b^2} y_0 & \text{if } y_0 < \left( \frac{b}{\sqrt{a^2 - b^2}} - r^+ \right) |x_0| \\ U(x_0, y_0) = \frac{b}{\sqrt{a^2 - b^2}} y_0 & \text{otherwise} \end{cases}$$

**Proof**

We just have to show that, whatever are  $a > b > c > 0$ ,  $J_1(x_0, y_0) = J_2(x_0, y_0)$  is reached by  $(x_0, y_0)$  such that:

$$t_{f_1}(y_0) < \frac{|x_0|}{c} < t_{f_2}(y_0).$$

- $J_1(x_0, y_0)$  cannot be strictly greater than  $J_2(x_0, y_0)$  when the condition  $t_{f_1}(y_0) = \frac{|x_0|}{c}$  is met. Otherwise,  $J_1$  and  $J_2$  being continuous with respect to  $y_0$ , there exists  $\epsilon > 0$  such that  $J_1(x_0, y_0 + \epsilon) > J_2(x_0, y_0 + \epsilon)$ , in a region where it has just been proved that the strategy 2 is maximizing.
- When the condition  $t_{f_2}(y_0) = \frac{|x_0|}{c}$  is met, the strategy 2 presents a maximizing trajectory without switching, which is obviously worse than the strategy 1 according to the trajectory slopes.

□

This example shows that in the region where the certainty equivalence principle does not hold, the value function does not follow a classical dynamic programming equation. Inside the area 2 of the state space,  $U$  depends only on  $y$ , but whatever are the choices of the players,  $y$  is strictly decreasing. So,  $U$  is never invariant along any of the trajectories compatible with the observation.

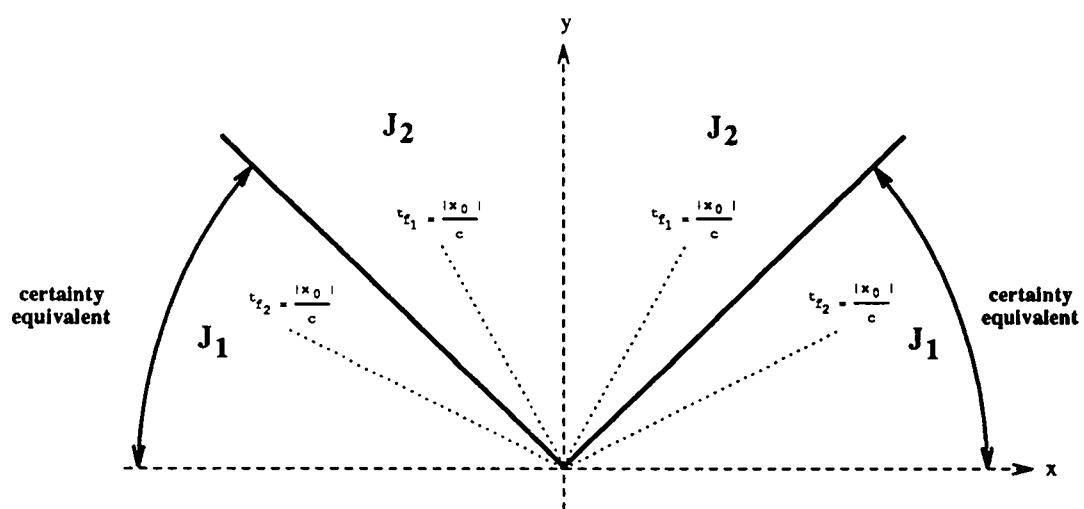


Figure 3: Partition of the domain of the value function

## Appendix

**Lemma 1** *If  $f$  is a function from  $\mathbb{R}$  to  $\mathbb{R}$  such that:*

$$\forall t \in \mathbb{R}, D^+ f(t) = \limsup_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} < 0$$

*then  $f$  is strictly decreasing.*

**Proof**

Let us choose  $a < b$ , then

$$\begin{aligned} \forall t \in [a, b], \exists \epsilon_t > 0 \mid & \forall t' \in (t - \epsilon_t, t + \epsilon_t), \\ & t' > t \Rightarrow f(t') < f(t) \\ & t' < t \Rightarrow f(t') > f(t). \end{aligned}$$

Otherwise the hypothesis of the lemma would not hold. The compact set  $[a, b]$  is so covered by open sets:

$$[a, b] \subset \bigcup_t (t - \epsilon_t, t + \epsilon_t)$$

from which one can extract a finite covering  $\{(t_i - \epsilon_{t_i}, t_i + \epsilon_{t_i})\}_{i \in I}$  (the  $t_i$  are ordered in an increasing sequence). Then there exists a sequence  $(a_i)$  such that:

$$\left\{ \begin{array}{ll} a_0 & = a \\ a_n & = b \\ a_i & \in (t_i, t_i + \epsilon_{t_i}) \cap (t_{i+1} - \epsilon_{t_{i+1}}, t_{i+1}) \\ a_i & < a_{i+1} \\ f(a_i) & < f(t_{i+1}) < f(a_{i+1}) \end{array} \right.$$

Thus  $f(a) < f(a_1) < \dots < f(a_{n-1}) < f(b)$ .

□

**Lemma 2** *If  $f$  is a function from  $\mathbb{R}$  to  $\mathbb{R}$  such that:*

$$\forall t \in \mathbb{R}, D^+ f(t) = \limsup_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} \leq 0$$

*then  $f$  is nonincreasing.*



**Proof**

If  $f$  strictly increases from  $a$  to  $b > a$ , let us define:

$$g(t) = f(t) - (t - a) \frac{f(b) - f(a)}{b - a}$$

over  $[a, b]$ . Then,

$$D^+g(t) = D^+f(t) - \frac{f(b) - f(a)}{b - a} < 0$$

and so, according to the previous lemma,  $g$  is strictly decreasing. But,  $g(a) = g(b) = f(a)$ , a contradiction.

□

## References

- [Ba90] T. Başar. Game theory and  $H^\infty$ -optimal control. In *4th international conference on Differential Games Approach*. Springer-Verlag, August 1990.
- [BB91] T. Başar and P. Bernhard.  *$H^\infty$ -Optimal control and Related Minimax Design Problems: A Dynamic Game Approach*. Birkhäuser, Boston, 1991.
- [Ber87] P. Bernhard. Differential games: Isaacs' equation. In Madan Singh, editor, *Encyclopedia of Systems and Control*. Pergamon Press, 1987.
- [Ber90a] P. Bernhard. A certainty equivalence principle and its applications to continuous-time, sampled data, and discrete time  $H^\infty$  optimal control. Technical Report 1347. INRIA, August 1990.
- [Ber90b] P. Bernhard. Variations sur un thème de danskin avec une coda sur un thème de von neumann. Technical Report 1238, INRIA Sophia-Antipolis, June 1990.
- [Dan67] J.M. Danskin. *The Theory of Max Min*. Springer, Berlin, 1967.
- [DBB93] G. Didinsky T. Başar and P. Bernhard. Structural properties of minimax controllers for a class of differential games arising in nonlinear  $H^\infty$ -control. *Syst. Contr. Lett.*, 1993. to appear.
- [Isa65] R. Isaacs. *Differential Games*. Wiley, 1965.
- [KS77] N. Krassovski et A. Soubbotine. *Jeux différentiels*. Editions MIR, 1977.



---

Unité de Recherche INRIA Sophia Antipolis  
2004, route des Lucioles - B.P. 93 - 06902 SOPHIA ANTIPOLIS Cedex (France)

Unité de Recherche INRIA Lorraine Technopôle de Nancy-Brabois - Campus Scientifique  
615, rue du Jardin Botanique - B.P. 101 - 54602 VILLERS LES NANCY Cedex (France)  
Unité de Recherche INRIA Rennes IRISA, Campus Universitaire de Beaulieu 35042 RENNES Cedex (France)  
Unité de Recherche INRIA Rhône-Alpes 46, avenue Félix Viallet - 38031 GRENOBLE Cedex (France)  
Unité de Recherche INRIA Rocquencourt Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

---

EDITEUR  
INRIA - Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

ISSN 0249 - 6399



★ R R - 2 8 1 9 ★